Molecular dynamics simulation with finite electric fields using Perturbed Neural Network Potentials

Kit Joll¹, Philipp Schienbein*^{1,2}, Kevin M. Rosso³, and Jochen Blumberger^{†1}

 Department of Physics and Astronomy and Thomas Young Centre, University College London, London, WC1E 6BT, United Kingdom
Department of Physics, Imperial College London, Exhibition Rd, South Kensington, London, SW7 2AZ, United Kingdom
Pacific Northwest National Laboratory, Richland, Washington 99354, United States

^{*}to whom correspondence should be addressed, email: p.schienbein@ucl.ac.uk †to whom correspondence should be addressed, email: j.blumberger@ucl.ac.uk

Abstract

The interaction of condensed phase systems with external electric fields is of major importance in myriad processes in nature and technology ranging from the field-directed motion of cells (galvanotaxis), to geochemistry and the formation of ice phases on planets, to field-directed chemical catalysis, to energy storage and conversion systems including supercapacitors, batteries and solar cells. Molecular simulation in the presence of electric fields would give important atomistic insight into these processes but applications of the most accurate methods such as ab-initio molecular dynamics are limited in scope by their computational expense. Here we introduce Perturbed Neural Network Potential Molecular Dynamics (PNNP MD) to push back the accessible time and length scales of such simulations at virtually no loss in accuracy when compared to ab-initio molecular dynamics. The total forces on the atoms are expressed in terms of the unperturbed potential energy surface represented by a standard neural network potential and a field-induced perturbation obtained from a series expansion of the field interaction truncated at first order. The latter is represented in terms of an equivariant graph neural network, trained on the atomic polar tensor. PNNP MD is shown to give excellent results for the dielectric relaxation dynamics, the dielectric constant and the field-dependent IR spectrum of liquid water when compared to ab-initio molecular dynamics or experiment, up to surprisingly high field strengths of about $0.2\,\mathrm{V\,\AA}^{-1}$. This is remarkable because, in contrast to most previous approaches, the two neural networks on which PNNL MD is based are exclusively trained on zero-field molecular configurations demonstrating that the networks not only interpolate but also reliably extrapolate the field response. PNNP MD is based on rigorous theory yet it is simple, general, modular, and systematically improvable allowing us to obtain atomistic insight into the interaction of a wide range of condensed phase systems with external electric fields.

Introduction

Electric fields are omnipresent in nature and technology. Their presence guides bumblebees to find nectar [1], and they are believed to cause a superionic ice VII phase on Venus [2]. They play a central role in myriad electronics and energy conversion devices including field effect transistors, (super-)capacitors, batteries and solar cells. In chemistry, electric fields can be used to steer selectivities in catalysis [3] and to control reactivities [4], whilst in physics they are used to accelerate particles close to the speed of light. The field strengths in these examples span an extraordinarily large range. Atmospheric fields in fair weather are on the order of 10^{-6} V Å^{-1} [5] and electric fields acting as floral cues can be as large as 10^{-5} V Å^{-1} [1]. In electrical components the field strength can vary significantly depending on the actual design and the envisaged application from about 10^{-10} to 10^{-3} V Å^{-1} . Particle accelerators typically operate at a field strength of up to $5 \times 10^{-3} \text{ V Å}^{-1}$. At charged electrodes field strengths on the order of 0.1 V Å^{-1} [6] can be found, which also marks the onset of chemical bond activation [7]. The making and breaking of chemical bonds may require more than 1 V Å^{-1} [8, 9].

Molecular dynamics (MD) or Monte Carlo (MC) simulations are the most suitable methods to understand and predict the interaction of liquids, liquid electrolytes and liquid/solid interfaces with external electric fields because they sample the correct equilibrium distribution of molecular configurations. Many finite electric field simulations [10] have been carried out with classical force fields [11, 12, 13] which typically give a good description for pure solvents but tend to perform poorly on more complex systems where charge transfer and polarization effects become important, e.g., electrode/electrolyte interfaces. Ab-initio MD (AIMD) simulations[14] solve the electronic structure of the system from first principles at every MD time step (usually at the density functional theory (DFT) level of theory) and thus give the most accurate description for such cases. Indeed, AIMD simulations with finite electric fields have been carried out on a number of systems[15], ranging from crystals [10] and pure liquid water[16, 17] to electrode/electrolyte interfaces[18, 19, 20, 21]. A serious disadvantage of AIMD simulations is that they are still computationally extremely demanding, with simulation times typically limited to a few 10 picoseconds and system sizes limited to a few 100 atoms depending on the density functional chosen. As such, AIMD simulations of important field-induced phenomena, such as dielectric relaxation, ionic conductivity[13], electric double layer formation[22] or capacitive charging are prohibitively expensive due to the large number of atoms required to faithfully model such processes.

The advent of machine learning (ML) has transformed the field of MD simulations. Training a ML potential (MLP) from a few hundred explicit electronic structure calculations[23, 24, 25], it is now possible to carry out MLMD simulation at AIMD accuracy on truly nanoscale systems over nanosecond time scales. Originally, these ML models were designed to calculate potential energies and forces only. Over the last years, a full bouquet of ML models have been published to predict dipole moments and polarizabilities [26, 27, 28, 29, 30, 31, 32], as well as other response properties including atomic polar tensors (APT)[33]. Following these developments, some ML models have been introduced which explicitly describe the interaction of a molecular system with external electric fields[34, 35, 36, 37, 38]. In most of these approaches the field-dependence of the potential energy surface is explicitly or implicitly part of the MLP [34, 35, 36, 38]. Whilst these approaches could be successfully applied to molecular systems and liquids, the electric field is an input parameter to the MLP and therefore training data for different field strengths is required.

Herein, we introduce a simple, complementary approach for MLMD with electric fields that does not require the generation of any training data with an electric fields. We start from a standard potential energy surface and account for the interaction with the electric field in a perturbative manner by a series expansion truncated at first order, noting that the first order term can be written in terms of the APT [39, 33]. Two MLPs are trained: one standard MLP for the unperturbed potential energy surface (here, a committee neural network potential (c-NNP) [23, 40, 25]) and one ML model for the APT (here, an equivariant graph neural network denoted "APTNN" [33]). Combined they form a "perturbed MLP" (here, "perturbed neural network potential" (PNNP)). The two networks have no knowledge of electric field effects (neither explicitly, nor implicitly) because they are both trained solely on data and atomic configurations obtained from zero-field calculations. The interaction with the field during MD simulation simulation is therefore entirely due to the first order term in the series expansion given by the product of the APT represented by the graph neural network and the external field.

A PNNP is based on rigorous physical principles and follows precisely the treatment of electric fields in quantum chemical calculations, where the total energy is usually expanded at zero field [41]. The accuracy can thus be systematically improved by adding higher order contributions, such as polarizabilities and hyperpolarizabilities, in ML representation. Since we do not introduce or modify ML models, the advantages and accuracies, but also the limitations of the two employed ML models are inherited. Importantly, a PNNP follows the spirit of the modern theory of polarization [42] because the APT relates to a *change* of polarization and is thus not affected by the multivaluedness of the polarization in periodic boundary conditions. This makes PNNP applicable to a broad range of condensed phase systems typically modelled under periodic boundary conditions, including solids, liquids, and interfaces.

In the following section we describe the PNNP approach in detail. After validation we apply the method to simulate the dielectric response of pure liquid water. We demonstrate reversible polarization and depolarization of liquid water as the field strength is stepped up and down. Then the dielectric relaxation dynamics is analysed in detail, followed by the calculation of the dielectric constant from the response of polarization with respect to the field strength resulting in excellent agreement with the experimental value. Moreover, we show that PNNP correctly predicts the field-induced red shift in the O-H stretching mode and the field-induced blue shift in the librational mode of liquid water, in very good agreement with results for ab-initio molecular dynamics. Perhaps the most surprising aspect of PNNP is that it predicts forces, energies and the dielectric response in excellent agreement with ab-initio data up to high electric fields, even though the PNNP was only trained on zero-field equilibrium structures. In *Discussion* we compare PNNP to previously introduced ML methods for calculation of molecular systems with electric fields and close our work in *Conclusions*.

Results

Perturbed Neural Network Potential (PNNP)

In our approach the interaction of the atomistic system with a homogeneous external electric field **E** is treated perturbatively via a series expansion truncated at first order in the field [41, 18, 20, 13]

$$\mathcal{H}_{\mathbf{E}}(\mathbf{r}^{N}, \mathbf{p}^{N}) = \mathcal{H}_{0}(\mathbf{r}^{N}, \mathbf{p}^{N}) - \mathbf{E} \cdot \mathbf{M}(\mathbf{r}^{N}), \qquad (1)$$

where $\mathcal{H}_0(\mathbf{r}^N, \mathbf{p}^N)$ is the total unperturbed Hamiltonian comprised of the kinetic energy of the N nuclei with momenta \mathbf{p}^N and the electronic potential energy depending on all nuclear positions \mathbf{r}^N , $\mathcal{H}_0(\mathbf{r}^N, \mathbf{p}^N) = E_{\text{kin}}(\mathbf{p}^N) + E_{\text{pot}}(\mathbf{r}^N)$, and $-\mathbf{E} \cdot \mathbf{M}(\mathbf{r}^N)$ is the perturbation induced by the electric field \mathbf{E} acting on the total dipole moment of the system at zero field $\mathbf{M}(\mathbf{r}^N)$. The truncation to first order in the field is expected to be accurate for the weak and medium strong electric fields investigated in this work, as will be demonstrated further below. The field dependence

of the dipole moment or, equivalently, higher order terms (polarizability, hyperpolarizability) may be added at stronger fields to account for the field-dependent perturbation of the electronic structure. [43] Applying Hamilton's equation of motion, we get for the force acting on atom i

$$F_{i\xi} = -\frac{\partial E_{\text{pot}}(\mathbf{r}^N)}{\partial r_{i\xi}} + \sum_{\zeta} \frac{\partial M_{\zeta}}{\partial r_{i\xi}} E_{\zeta}, \qquad (2)$$

where $\zeta = x, y, z$ and $\xi = x, y, z$ represent the three Cartesian coordinates and E_{ζ} is the ζ -component of \mathbf{E} . The first term is the force on the nuclei in the absence of an electric field and the second term is the field-induced contribution which can be written in terms of the transpose of the APT of atom i, $\mathcal{P}_{i}[33, 39]$ (not to be confused with the polarization \mathbf{P} , Eq. 6), with elements

$$\frac{\partial M_{\zeta}}{\partial r_{i\xi}} \equiv [\mathcal{P}_i^{\mathrm{T}}]_{\xi\zeta}.\tag{3}$$

In our approach we train ML models, one for the potential energy $(E_{\text{pot}}(\mathbf{r}^N))$ and one for the APT (\mathcal{P}_i) and use the corresponding forces, Eq. 2, to carry out MD simulations in the presence of an external electric field. We emphasize that both quantities are trained without an external field present, in contrast to the schemes suggested before [34, 35, 36, 38]. We use a committee [40] of 2nd generation high-dimensional Neural Network potentials [23, 25] (c-NNP) to model $E_{\text{pot}}(\mathbf{r}^N)$ and a E(3)-equivariant graph neural network to model the APT (APTNN) as recently introduced by one of us [33]. For details of the force implementation we refer to section Methods. In the following the combined c-NNP and APTNN model for the electronic potential energy including the field term, $E_{\text{pot}}(\mathbf{r}^N) - \mathbf{E} \cdot \mathbf{M}(\mathbf{r}^N)$, is simply referred to as "perturbed neural network potential" (PNNP).

PNNP MD simulations give access a number of important dielectric properties of solids, liquids and ionic solutions. The time derivative of the total dipole moment can be obtained by summing all APTs multiplied by the respective velocities of the nuclei, $v_{i\xi}[33]$

$$\dot{\mathbf{M}} \equiv \frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = \sum_{i,\xi} \frac{\partial \mathbf{M}}{\partial r_{i\xi}} v_{i\xi} . \tag{4}$$

Thus, field-dependent IR spectra can be readily obtained from the autocorrelation function of $\dot{\mathbf{M}}$ sampled along PNNP trajectories (see Eq. 9 below). Time-integration gives the total dipole moment of the cell,

$$\mathbf{M}(t) = \mathbf{M}(t_0) + \int_{t_0}^t \mathrm{d}t' \dot{\mathbf{M}}(t'), \tag{5}$$

and the polarization \mathbf{P} ,

$$\mathbf{P(t)} = \frac{\mathbf{M(t)}}{V},\tag{6}$$

where V the volume of the simulation box. In passing we refer to the modern theory of polarization in solids[43], where $\dot{\mathbf{M}}(t)/V$ is the transient current density that, when integrated over time gives the itinerant dipole moment. The polarization in Eq. 6 is of major importance allowing for the calculation of relevant dielectric properties including the dielectric constant (or "relative permittivity", see Eq. 7 below), capacitance and ionic conductivity.

Validation of PNNP

We demonstrate our approach by simulation of pure liquid water at room temperature for different electric field strengths ranging from 0.0026 to $0.2057 \,\mathrm{V\, \mathring{A}}^{-1}$ (corresponding to 5×10^{-5} and

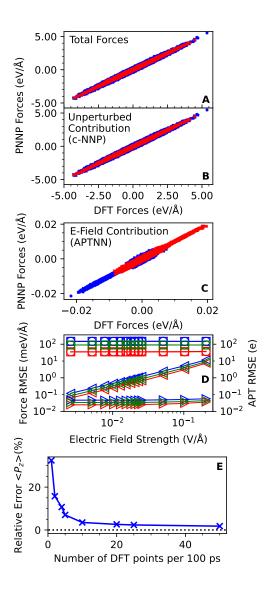


Figure 1: Error metrics of the trained PNNP at an exemplary field strength of $0.0129\,\mathrm{V\,\mathring{A}}^{-1}$. Scatter plots are presented comparing the predicted PNNP force with the corresponding reference DFT forces for the total PNNP force (A) and its individual components, namely the unperturbed force given by the c-NNP (B) and the field induced force obtained by the APTNN (C). Forces acting on O and H atoms are colored in blue and red, respectively. Mind the smaller scale in panel C since the field-induced forces are about two magnitudes smaller than the unperturbed forces. Panel D shows the overall RMSE as a function of the strength of the applied electric field, where the squares, circles, and left-facing triangles depict the force RMSE of the total PNNP force, the zero-field force, and the field-induced force, respectively. The right-facing triangles indicate the RMSE of the APT in units of elementary charge (right y axis). Mind the logarithmic scale and that the total PNNP force (squares) and the zero-field force (circles) largely overlaps. In E the relative error of the mean polarization in the direction of the field (P_z) obtained by integrating the total dipole moment time derivative (Eq. 5) as a function of the sampling frequency of explicit DFT points, see text, is plotted.

 4×10^{-3} atomic units, respectively). Recall, that the latter field strength is well in the regime of chemical bond activation [7]. Details on the training of the c-NNP and APTNN for liquid water are given in section Methods. Running a PNNP trajectory at an intermediate field strength of $0.0129 \,\mathrm{V\,\AA}^{-1}$, the drift in the conserved total energy is $-2.2 \times 10^{-9} \,\mathrm{Hartree\,atom}^{-1}\,\mathrm{ps}^{-1}$ and the mean magnitude of the center of mass momentum is 8.4×10^{-9} atomic units, see Fig. S1 A and B, respectively. These values are representative for all performed PNNP simulation, i.e. all applied electric field strengths, and they are also typical for simulations without field attesting to the correct force implementation. Next we assess the quality of PNNP by validating the predicted forces along the trajectories against an independent test set unknown to the PNNP. At each field strength, 100 configurations are equidistantly extracted from 100 ps PNNP trajectories. Then the total force on each nuclei is computed using DFT reference calculations at the same field strength and compared to the total force from PNNP (data shown in Fig. 1A for an exemplary field strength of $0.0129 \,\mathrm{V\,\AA}^{-1}$). We then decompose the DFT reference force into unperturbed and electric field-induced contributions by performing an additional DFT calculation on the same configurations, but without the external electric field. The difference between the forces with and without the field isolates the field-induced forces. This allows us to separately validate the unperturbed c-NNP force contribution, first term on the right-hand side of Eq. 2, with the unperturbed DFT contribution (Fig. 1B) and the electric field-induced APTNN contribution, second term on the right-hand side of Eq. 2, with the field-induced DFT contribution (Fig. 1C).

We find excellent correlation for PNNP and DFT total forces and its contributions. The root-mean-squared-error (RMSE) in the atomic forces as a function of field strength are shown in Fig. 1D. The total force RMSEs are on the order of 90 meV Å $^{-1}$ and thus in line with previously published HD-NNP and c-NNPs on various systems [44, 45, 40, 46]. Remarkably, we find that the RMSE for the unperturbed contribution remains almost constant at about 90 meV Å $^{-1}$ across all simulated field strengths up to $0.2057\,\mathrm{V\,\mathring{A}}^{-1}$, although the training set is only composed of unperturbed, i.e., zero-field configurations and does not contain any polarized water samples. The field-induced force contributions are about two orders of magnitude smaller than the unperturbed contribution and so are their RMSEs. Notably, the force RMSE increases only in proportion with the field strength from $8.6\times10^{-2}\,\mathrm{meV\,\mathring{A}}^{-1}$ at $0.0026\,\mathrm{V\,\mathring{A}}^{-1}$ to $9.05\,\mathrm{meV\,\mathring{A}}^{-1}$ at the largest simulated field of $0.2057\,\mathrm{V\,\mathring{A}}^{-1}$. The reason for this is that the RMSE of the APT itself (Fig. 1D, left y axis) remains close to constant across all simulated field strengths. Due to error propagation (see Eq. 2) that constant RMSE is amplified by the magnitude of the respective field strength. Note that even at the highest field strength investigated, the force RMSE of the field contribution is about one magnitude smaller than the RMSE of the c-NNP.

Lastly, we validate the calculation of the total dipole moment along the PNNP MD trajectories as obtained by integration of the time derivative of the dipole moment according to Eq. 5. We find that the dipole moment tracks the reference dipole moments obtained from explicit DFT calculations very well for about 10-100 ps depending on the field strength (Fig. S1C). However, at longer times deviations become large due to accumulation of errors when integrating over the finite time steps, even though $\dot{\mathbf{M}}(t)$ is accurately reproduced. This problem is addressed here by calculating reference DFT dipole moments along PNNP MD trajectories in periodic intervals and integrating the time derivative of the dipole moment obtained from APTNN only from one DFT reference value to the next. Using this integration procedure we calculated the mean dipole moment averaged over the 100 ps PNNP MD trajectories, $\langle \mathbf{M}(t) \rangle$, as a function of the time interval between two DFT reference values. The relative errors with respect to the reference DFT mean dipole moments averaged over the same 100 ps PNNP MD trajectories (calculated using a sampling frequency of 1 ps⁻¹) are shown in Fig. 1E. We find that the error decreases rapidly, to below 5% for a time interval between two DFT reference dipole moments of 10 ps,

for all field strengths. Hence, in practice only a limited number of additional DFT calculations are necessary to accurately calculate the dipole moment along the PNNP MD trajectories at a resolution that is only limited by the MD time step.

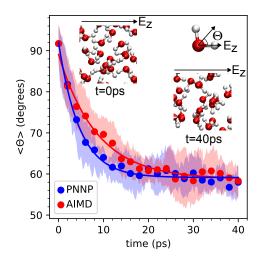


Figure 2: Orientational relaxation of water molecules in response to an applied electric field being switched on at t=0. The time evolution of the average water orientation angle $\langle\Theta\rangle$ comparing AIMD (red) and PNNP simulations (blue). The angle Θ is defined for each water molecule as the angle between the bisector of the two intramolecular OH bond vectors and the applied electric field vector E_z (as illustrated by the inset in the top right corner). The average $\langle\Theta\rangle$ is obtained by averaging Θ over all water molecules in a given configuration. The solid circles depict the average of $\langle\Theta\rangle$, while the shaded area indicates its corresponding standard deviation. The solid line represents an exponential fit to the solid circles. Note that 90° corresponds to a random orientation on average.

Orientational relaxation dynamics

Having validated the forces and polarization against DFT reference data, we now apply the PNNP to simulate the relaxation of water orientation in response to interaction with an electric field. We extract 10 independent configurations from a sample of liquid water equilibrated for 1 ns with field-free c-NNP MD and use them as starting configurations for 40 ps PNNP MD simulations, each run at a field strength of $0.0257 \,\mathrm{V\,\mathring{A}}^{-1}$. Additionally, we take the same initial configurations and perform explicit AIMD simulations at the same field strength. The orientation of the water molecules at a given time t is described by the angle between the direction of the applied electric field and the bisector between the two intramolecular OH bond vectors averaged over all water molecules, $\langle \Theta \rangle(t)$. The results are shown in Fig. 2. At t=0 the initial average orientation is 90° corresponding to the expectation value of the randomly oriented water molecules at equilibrium. When the field is switched on at t=0 the relaxation dynamics obtained from PNNP MD is in excellent agreement with the results from AIMD for the first 5 ps and at longer times > 20 ps, both converging to a final average angle of 58°. Small deviations can be observed between 5-20 ps where PNNP MD gives a slightly faster exponential decay (time constant $\tau=4.66$ ps⁻¹,

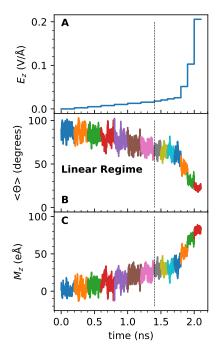


Figure 3: Progression of the performed field sweep MLMD simulation as a function of time. The magnitude of the applied electric field (E_z) is presented in panel A, while the panels B and C show the average water orientation angle with respect the the field $(\Theta, \text{same as in Fig. 2})$ and the total dipole moment (M_z) in the direction of the field, respectively. The vertical dashed line at 1.4 ns indicates when the electric field strength exceeds $0.0154 \,\mathrm{V\, \mathring{A}}^{-1}$ for the first time, thereby reaching the non-linear response regime. Before 1.4 ns each field strength was kept constant for 200 ps, afterwards it was kept constant for 100 ps.

 $R^2 = 0.98$) than the data from AIMD ($\tau = 8.53 \text{ ps}^{-1}$, $R^2 = 0.99$). However, since the error bars for the two methods overlap we ascribe this difference to statistical uncertainty.

Field Sweep

In the following we vary the strength of the field in time and monitor the structural and dielectric response of the water sample as obtained from PNNP simulations. The results are presented in Fig. 3. The temporal profile of the applied electric field strength is shown in panel A, the average water orientation $\langle\Theta\rangle(t)$ in panel B and the total dipole moment in the direction of the applied electric field, $M_z(t)$, in panel C. We obtain an approximately linear response in the average orientation and a concomitant linear increase in the dipole moment as the field is stepped up from 0 to $0.0154 \,\mathrm{V\, \mathring{A}}^{-1}$ (in increments of $0.0026 \,\mathrm{V\, \mathring{A}}^{-1}$). This is in line albeit somewhat lower than previously published estimates for the upper bound of the linear regime, $0.03 \,\mathrm{to}$ $0.07 \,\mathrm{V\, \mathring{A}}^{-1}$ [15, 47]. At larger field strengths the dielectric response becomes weaker indicating

that the non-linear regime is reached. At $0.2057\,\mathrm{V\,\mathring{A}}^{-1}$ a strong orientational alignment along the field direction is observed, $\langle\Theta\rangle=20^\circ$. Notice that at these high field strengths the PNNP still predicts reasonably accurate forces (see Fig. 1D) despite the absence of electronic polarization terms and any field-dependent training data. At $0.4114\,\mathrm{V\,\mathring{A}}^{-1}$, we observe water splitting into a proton and a hydroxide ion. It is well known that chemical bond activation starts to occur at about $0.1\,\mathrm{V\,\mathring{A}}^{-1}$ [7] and therefore this observation is not surprising. While we could in principle further train the c-NNP to also include proton and hydroxide species, and thereby test even larger field strengths, accurate description of these species would require explicit inclusion of nuclear quantum effects [48]. An additional set of simulations are run where the field strength is reversed from the value at the end of the linear regime, $0.0154\,\mathrm{V\,\mathring{A}}^{-1}$, down to 0 in increments of $0.0026\,\mathrm{V\,\mathring{A}}^{-1}$. That sweep is detailed in Fig. S2. We obtained very similar mean orientations and dipole moment as for the forward sweep demonstrating that the sample can be reversibly polarized and depolarized.

Dielectric constant

The time series of dipole moments in Fig. 3C are averaged individually for each applied electric field strength to obtain the polarization of the water sample along the field direction, P_z (Eq. 6), as a function of the field strength. The results are shown in Fig. 4 for the forward and backward sweeps (data in blue and red, respectively). The data points in the linear regime are shown magnified in the inset of Fig. 4. They are fit to straight lines and the slopes used to obtain the static dielectric constant, ϵ_r , according to Eq. 7,

$$\epsilon_r = 1 + \frac{1}{\epsilon_0} \frac{\partial \langle P_z \rangle}{\partial E_z} \,, \tag{7}$$

where ϵ_0 is the vacuum permittivity. We obtain values $\epsilon_r = 77.9 \pm 2.7 \ (R^2 = 0.99)$, $80.6 \pm 2.7 \ (R^2 = 0.96)$ for forward and backward sweep and $\epsilon_r = 79.3 \pm 2.2 \ (R^2 = 0.99)$, data in green) from a linear fit to the weighted average data for forward and backward sweep, in excellent agreement with the experimental value of $\epsilon_r = 78.4 \ [49]$. Similar values would be obtained using only one data point due to the excellent linear correlation. Remarkably, the dielectric constant was converged after only about 175ps of simulation time per applied electric field (see Fig. S4), similar to what was previously reported for classical MD simulations employing finite field Hamiltonians[20]. This is an order of magnitude less simulation time than what is required to calculate the dielectric constant from the fluctuations of the polarization [50, 51, 52, 20],

$$\epsilon_r = \epsilon_{\infty} + \frac{1}{3\epsilon_0 k_{\rm B} V T} \left(\langle \mathbf{M}^2 \rangle - \langle \mathbf{M} \rangle^2 \right) ,$$
 (8)

where the total dipole moment, \mathbf{M} , is sampled at zero electric field and $\epsilon_{\infty} = 1.72$ [53] is the optical dielectric constant for liquid water. Indeed, sampling the total dipole moment \mathbf{M} along zero field c-NNP trajectories we obtain a converged value $\epsilon_r = 90.6 \pm 1.7$ only after 3 ns (see Fig. S3) similarly as in previous simulations [54]. The reason for the difference in the dielectric constant from the field sweep and from zero field simulation is not known but could be due to a number of reasons. The fluctuations of the polarization are likely to be more sensitive on simulation details than the mean values, e.g., thermostat used, finite system size and remaining inaccuracies of the PNNP. Moreover, the approximations made to derive the dielectric constant in terms of the polarization fluctuations from its basic definition Eq. 7 [55] could also contribute to the difference.

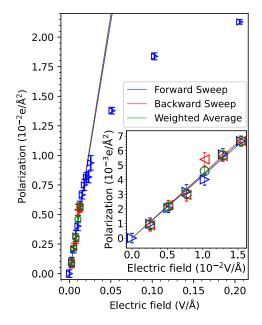


Figure 4: Average polarization as a function of the applied electric field strength. The blue right triangles depict the average polarization sampled from the field sweep PNNP MD simulation in Fig. 3 going from small to large fields ("forward-sweep"). Additionally, a backward sweep was performed (red left triangles) going from $0.0154\,\mathrm{V\,\mathring{A}^{-1}}$ (the end of the linear regime) back to zero field, see Fig. S2. The green open circles mark the weighted average of the forward and backward sweep. The solid straight lines depict the linear fit to the respective data set (blue: forward sweep, red: backward sweep, green: weighted average), considering the linear response regime only ($E \leq 0.0154\,\mathrm{V\,\mathring{A}^{-1}}$), from which the dielectric constant is calculated (see Eq. 7). The inset contains an amplification of the linear response regime, showing the very same data and using the same color code; mind the smaller scales on both axes.

Field-dependent IR spectra

The APTs and velocities along the PNNP trajectories give direct access to the time derivative of the dipole moment, $\dot{\mathbf{M}}$, according to Eq. 4. The autocorrelation function of $\dot{\mathbf{M}}$ gives the frequency (ω)-dependent Beer-Lambert absorption coefficient of IR spectroscopy, $\alpha(\omega)$,

$$\alpha(\omega)n(\omega) = \frac{\pi}{3Vc\epsilon_0 k_{\rm B}T} \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \, e^{-i\omega t} \left\langle \dot{\mathbf{M}}(0)\dot{\mathbf{M}}(t) \right\rangle,\tag{9}$$

where $n(\omega)$ is the frequency-dependent refractive index, c is the speed of light in vacuum, $k_{\rm B}$ the Boltzmann constant and T the temperature. In Fig. 5 we present the calculated $\alpha(\omega)n(\omega)$ as obtained from c-NNP simulations at zero field and from PNNP trajectories at different field strengths in the non-linear response regime. The experimental IR spectrum at zero-field [56] is very well reproduced using the RPBE-D3 functional, as reported previously [33]. The spectrum remains remarkably insensitive to the presence of an electric field in the linear regime and beyond, up to about $0.0514 \, {\rm V \, \mathring{A}}^{-1}$. For larger fields we observe a systematic red-shift of the intramolecular

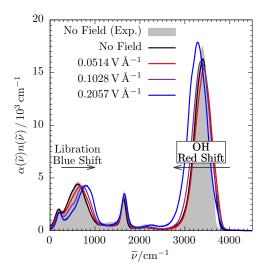


Figure 5: Field-dependent Beer-Lambert absorption coefficients of liquid water. The Beer-Lambert absorption coefficient times the frequency dependent refractive index $\alpha(\widetilde{\nu})n(\widetilde{\nu})$ of room temperature liquid water is presented as a function of the applied electric field strength. The black solid line depicts the spectrum of water without the presence of any field, and the red, purple, and blue curves depict the corresponding spectra for field strengths of 0.0514, 0.1028, and 0.2057 V Å⁻¹, respectively. The experimental spectrum at zero field is taken from Ref. [56] and shown for reference as shaded grey area.

OH stretching vibration at around 3500 cm⁻¹, by approximately 100 cm⁻¹ at 0.2057 V Å⁻¹. A red-shift of the OH stretch is generally ascribed to formation of stronger intermolecular hydrogen bonds [57, 58, 59], here induced by the external electric field. This notion is further supported by the systematic blue shift of the librational band at around 3500 cm⁻¹ by approximately 200 cm⁻¹ at 0.2057 V Å⁻¹ suggesting that the interaction with the field leads to a stiffening of the potential for rotational motion of the water molecules [60]. This is in line with a previous study describing water at these field strengths to be more "ice-like" [61]. The trends in the IR spectra agree very well with previously published results obtained from other approaches that differ in many aspects from our method, AIMD simulations with external electric fields [61] and MLMD simulations using the FIREANN-wF model [38].

Discussion

In this work we extend MLMD simulations to include the interaction with an external electric field by adding the field induced perturbation up to first-order to the unperturbed Hamiltonian Eq. 1. The resulting force equation (Eq. 2) allows us to separately calculate the unperturbed forces from a standard MLP (here: a c-NNP) and the field-induced forces from the APTNN. The force calculation therefore rigorously follows from the laws of electrostatics without any additional approximations. This has several advantages: First, the approach is modular because the electric field contribution is independent from the employed unperturbed potential energy surface. As such, the APTNN can be coupled with any MLP (not only NNPs) or even with force fields to include the interaction with an external electric field. The level of theory of the reference

electronic structure calculations for the training of APTNN and MLP can be chosen differently in accord with the accuracy requirements for unperturbed and perturbed potentials. Moreover, the modular approach allows one to assess the accuracy of each component separately. Second, APTs are well defined for any atomistic system and thus APTNNs can be trained "out of the box", without requiring any conceptual tailoring or adjustment or the use of arbitrarily defined proxies (e.g. atomic charges, molecular dipole moments, or similar)[33]. Importantly, APTs are uniquely defined when using periodic boundary conditions as they quantify a *change* of the total dipole moment or polarization. This is in contrast to the total dipole moment, which is multi-valued in periodic boundary conditions [42]. Third, the method can be systematically improved by including higher order terms to the perturbation, e.g. polarizability and hyperpolarizability [41]. This would allow for more accurate simulations at even larger field strengths as investigated herein. Through the inclusion of the next higher term in the multipole expansion [62], one could potentially also describe electric field gradients. Finally, the approach does not require training as a function of the applied electric field, even up to rather strong fields at the onset of chemical bond activation.

These features make our method distinct from other approaches recently introduced to model the interaction with external electric fields in ML simulations. Field-dependent training data is usually required because the electric field is directly embedded into a single MLP [34, 35, 36, 38]. For example, to perform finite-field simulations on liquid water, the FIREANN-wF model trained atomistic forces in the presence of an electric field [38]. The model was then successfully used to calculate response functions, such as dipole moments and polarizabilities, and from these quantities field-dependent vibrational spectra of liquid water [38]. However, the field-induced contribution to the total energy and forces is thus only "learned implicitly in the force-only training" [38]. In contrast to our method, this approach is not modular and it is not obvious how it can be systematically improved. Earlier approaches [34, 35] rely on the definition of proxies, such as fictitious partial charges [34] or atomistic dipole moments [35]. It also does not depend on the definition of molecular reference frames, which is e.g. required to train the position of Wannier centers [36, 63]. The APTNN in turn, which is entirely responsible for the field-induced atomic forces, trains on a well-defined physical observable [39, 33] such that it is guaranteed to be meaningful independent of the system of interest. Finally, our method exclusively trains on unperturbed, zero-field data, the electric field dependence is therefore offloaded to the MD sampling by Eq. 2. As such, the electric field is not an input parameter of the ML model.

We performed PNNP simulations at several different field strengths and compared our results against reference AIMD calculations and experimental literature data. In general, we found very good to excellent agreement for all properties investigated. We demonstrated reversible orientational polarization and depolarization of the water molecules as the electric field is stepped up and down, with orientational relaxation times that are in good agreement with AIMD data (Fig. 2). This permitted successful calculation of the relative permittivity, ϵ_r , on simulation time scales that are an order of magnitude shorter than in more standard approaches that suffer from the very slow convergence of the polarization fluctuations at zero field. We obtained a value $\epsilon_r = 79.3 \pm 2.2$ in excellent agreement with experiment, $\epsilon_r = 78.41$ at 298.14 K and 1 atm [49] (Fig. 4). By definition (Eq. 7), the dielectric constant is a direct measure of a system's response to an applied electric field. The observed, almost quantitative, agreement with previous simulations and the experiment therefore validates our approach in the linear response regime below $0.0154 \,\mathrm{V\, \mathring{A}^{-1}}$. With respect to the non-linear regime, we calculated IR spectra and compared them with previous explicit AIMD simulations [61]. The systematic peak shifts in the IR spectrum for libration and intramolecular OH stretch vibration at high electric fields match that reference data very well, thereby validating our approach in the non-linear response regime as well.

The almost quantitative agreement obtained for the dielectric constant is remarkable, because one generally cannot expect perfect agreement between electronic structure calculations at the level of DFT and experiments. Moreover, nuclear quantum effects [48] were ignored in our simulations. The RPBE-D3 functional in combination with classical MD simulation is known to describe water, aqueous solutions and interfaces over a wide range of temperatures and pressures very well [60, 64, 65, 33, 54, 66, 67, 68]. With the help of PNNP simulations we could show that the excellent performance of the RPBE-D3 functional in describing radial distribution functions, self-diffusion, orientational relaxation processes, and vibrational spectra of liquid water with respect to experimental data, extends to its dielectric response properties. This is in part because remaining deficiencies of this functional tend to be effectively compensated by missing nuclear quantum effects. Indeed, our computed IR spectrum at zero-field is in almost quantitative agreement with experimental data. Moreover, our computed peak shifts in the IR spectra in the presence of an applied electric field are in almost quantitative agreement with the results reported in the MLMD simulations of Ref. [38]. Therein a more accurate and computationally expensive revPBE0-D3 hybrid functional was employed in combination with path integral MD simulations [38] to explicitly account for nuclear quantum effects.

We emphasize that in our method the machine learning models (c-NNP, APTNN) are exclusively trained on unperturbed, zero-field configurations and properties. Polarized water structures were not included in the training. Thus it is highly remarkable that the trained c-NNP and APTNN faithfully model "unseen" polarized structures up to field strengths of about $0.21 \,\mathrm{V\,\mathring{A}}^{-1}$. It also shows that inclusion of higher order terms in the perturbation series expansion, i.e. polarizabilities or hyperpolarizabilities, is not necessary at least up to this field strength, as far as the accuracy of nuclear forces in PNNP MD simulations is concerned. At a field strength of about $0.41 \,\mathrm{V\,\mathring{A}}^{-1}$ we observed the onset of water dissociation into $\mathrm{H^+}$ and $\mathrm{OH^-}$, causing the c-NNP to extrapolate and, ultimately, to fail. Our simple approach therefore covers a quite large regime of electric field strengths, from smaller fields in the linear regime up to chemical bond activation around $0.1 \,\mathrm{V\,\mathring{A}}^{-1}[7]$.

Conclusions

In conclusion, we have implemented a ML methodology to run MD simulations in the presence of an external homogeneous electric field. Key to this development is the APT, which is trained by a ML model and used to compute the field induced forces. The latter are combined with the forces obtained from a standard MLP describing the interatomic potential at zero-field. The method is rigorous, approximation-free, and does not require the definition of artificial proxies as target values for training, because the atomic polar tensor is a well-defined physical observable, even in periodic boundary conditions. By validating our approach against explicit AIMD simulations and experimental and computational data from the literature, we demonstrate that it accurately models liquid water in external homogeneous electric fields up to field strengths of about $0.21 \, \text{V Å}^{-1}$. This is deep in the non-linear response regime and at the onset for chemical bond activation. This is possible by training the MLP on zero-field configurations only. We expect that our method will be useful for ab-initio-level calculation of other field-dependent properties of condensed phase systems not covered in this work including ionic conductivity and capacitance.

Methods

Implementation of PNNP

The unperturbed potential energy, $E_{\rm pot}$, and the corresponding forces, first term on the right-hand side of Eq. 2 are modelled by a committee of 2nd generation high-dimensional Neural Network potentials, [23, 25] as recently implemented in the cp2k software package. [40] The field dependent force contribution related to the APT, second term on the right-hand side of Eq. 2, is modelled by an APTNN, which is based on the E(3)-equivariant graph neural network e3nn[69] based on the PyTorch library [70]. A new force evaluation environment was added to cp2k linking the fortran-based cp2k and the python/C++ based PyTorch in a client/server approach. Inspired by the i-PI implementation [71] cp2k launches and connects to a python server that waits to receive configurations and sends back APTs predicted by the APTNN model. The APTs are then used in cp2k, to evaluate the corresponding force contributions. They are added to the c-NNP forces using the built-in mixing force environment to obtain the total forces for propagation of the atoms using the velocity-Verlet algorithm.

Training of c-NNP and APTNN

The committee members of the c-NNP were trained using n2p2[72]. The parameters for the neural network were taken from a previous MLP study on liquid water[54], in conjunction with generic symmetry functions[40]. The c-NNP is trained on the energies and forces obtained from DFT calculations at the level of RPBE-D3 as detailed further below using the active learning procedure reported in Ref. [40]. With regard to the APT, we use the recently published APTNN developed by one us[33] containing 10,368 individual APTs in its training set at the level of RPBE-D3 [33]. This network has previously been benchmarked, and was shown to accurately reproduce the IR spectrum of liquid water under ambient conditions.

Reference DFT calculations

Electronic structure calculations for generation of reference training and test data were performed using $\mathtt{cp2k}[73]$ version 2023.1 and the quickstep module [74]. The RPBE functional was evaluated by the libxc package [75] and supplemented by D3 dispersion corrections [76]. We employed a mixed Gaussian orbital/plane wave (GPW) basis set [77] with a plane wave cutoff of 600 Ry and a relative cutoff of 20 Ry, and the Gaussian orbitals are constructed using the triple- ζ quality TZV2P basis set [78] including polarization functions as employed previously [66, 64, 65, 60]. Core electrons are described by norm-conserving Goedecker-Teter-Hutter (GTH) pseudopotentials [79, 80]. Homogeneous electric fields were treated by the approach introduced by Umari and Pasquarello [81] as implemented in cp2k. Dipole moments of the simulation cell (M) were obtained from the DFT calculations via maximally localized Wannier functions [14].

Charge conservation

Charge conservation is enforced by the acoustic sum rule

$$\sum_{i} \frac{\partial \mu_{\alpha}}{\partial R_{\beta}^{i}} = 0 \tag{10}$$

for the atomic polar tensors [82, 83, 84]. This is done by calculating the sum of every component divided by the number of atoms and distributing any excess evenly across all atoms in the system.

Note that this correction is akin to correcting atomic charges, as it has previously been done in the literature, distributing the excess charge evenly across the total system [85, 86, 87].

Simulation details

All calculations including training, testing and production runs were carried out for a 128 water molecule box of length 15.6627 Å employing periodic boundary conditions (density $0.996 \,\mathrm{kg} \,\mathrm{L}^{-1}$). All finite field PNNP MD simulations were carried out in the NVT ensemble using a CSVR thermostat with a time constant of 1 ps [88] and a MD integration time step of 1 fs. Zero field c-NNP MD simulations were carried out in the NVT ensemble using a Nose-Hoover thermostat[89, 90] and a time step of 0.5 fs.

To illustrate the reversability of the electric field within a single PNNP MD simulation, we performed a field sweep starting at zero field up to a field strength of $0.0154\,\mathrm{V\,\mathring{A}}^{-1}$ (the end of the linear regime). Therein, the field strength was increased every 200 ps by $0.0026\,\mathrm{V\,\mathring{A}}^{-1}$. Additionally, we then performed a backward sweep back to zero field by decreasing the applied electric field strength by $0.0026\,\mathrm{V\,\mathring{A}}^{-1}$ at every 200 ps as illustrated in Fig. S2 in the SI. The average water orientation angle $\langle\Theta\rangle$ (see main text) and the polarization in the direction of the field are plotted as well. From the slope of the polarization as a function of the applied field strength the dielectric constant can be estimated (see Eq. 7 in the main text). The field sweep PNNP MD can be used to calculate the dielectric constant twice, namely from the forward and from the backward sweep.

To calculate the Beer-Lambert absorption coefficients in Fig. 5, we follow a previous simulation protocol [60]: For each electric field strength (0.0514, 0.1028, and 0.2057 V Å $^{-1}$) 20 independent configurations were sampled from the field sweep PNNP simulation in Fig. 3. For each of these configurations short (20 ps) PNNP simulations in the NVE ensemble were conducted. The spectra are calculated for each of the 20 simulations independently and their average is shown in Fig. 5.

Convergence of the dielectric constant

The dielectric constant can either be calculated from the variance of the total dipole moment without any electric field present (Eq. 8) or from the response of the system to an applied electric field (Eq. 7). We first assess the convergence of the dielectric constant using the variance at zero field as shown in Fig. S3 in the SI. We observe that the variance is converged after simulating for about 3 ns in total (panel C), while sampling snapshots from the simulation every pico second (panel B). We also assess the convergence of the dielectric constant as a function of time using the change in polarization caused by an applied electric field (see Eq. 7 in the main text). To do so, we separately compute the mean polarization of the forward and backward field sweep up to time t and then perform a linear fit in full analogy to Fig. 4 yielding a dielectric constant for the forward and the backward sweep, respectively. The obtained dielectric constant as a function of time is presented in Fig. S4. We find that after about 175 ps, the value of the dielectric constant is converged. Each applied external field strength should thus be simulated for about 150 ps. Herein we run the simulation at each field strength for 200 ps (see above), taking equilibration to the new field strength into account. Due to the almost perfect linear fit of the polarization with respect to the applied field strength (see Fig. 4), only a few field simulations are practically required and thus the computational cost is thus significantly lower compared to converging the variance of the dipole moment at zero field (see Fig. S3).

Data Availability

The full data for this study totals more than 1 TB, so it is in cold storage accessible by the corresponding authors and available upon reasonable request.

Code Availability

All field simulations were performed using the cp2k software package, that has been customised in-house and will be made available in a GitHub repository linked in the supporting information. The atomic polar tensor neural network and all scripts used to train and predict atomic polar tensors is publicly available on: https://github.com/pschienbein/AtomicPolarTensor.

References

- [1] D. Clarke, H. Whitney, G. Sutton, and D. Robert. "Detection and Learning of Floral Electric Fields by Bumblebees". *Science* 340 (2013), pp. 66–69.
- Z. Futera, J. S. Tse, and N. J. English. "Possibility of realizing superionic ice VII in external electric fields of planetary bodies". Sci. Adv. 6 (2020), eaaz2915.
- [3] P. Besalú-Sala, M. Solà, J. M. Luis, and M. Torrent-Sucarrat. "Fast and Simple Evaluation of the Catalysis and Selectivity Induced by External Electric Fields". ACS Catal. 11 (2021), pp. 14467–14479.
- [4] S. Shaik, D. Danovich, J. Joy, Z. Wang, and T. Stuyver. "Electric-Field Mediated Chemistry: Uncovering and Exploiting the Potential of (Oriented) Electric Fields to Exert Chemical Catalysis and Reaction Control". J. Am. Chem. Soc. 142 (2020), pp. 12551–12562.
- [5] M. J. Rycroft, S. Israelsson, and C. Price. "The global atmospheric electric circuit, solar activity and climate change". J. Atmos. Sol. Terr. Phys. 62 (2000), pp. 1563–1576.
- [6] M. F. Toney, J. N. Howard, J. Richer, G. L. Borges, J. G. Gordon, O. R. Melroy, D. G. Wiesler, D. Yee, and L. B. Sorensen. "Voltage-dependent ordering of water molecules at an electrode-electrolyte interface". *Nature* 368 (1994), pp. 444–446.
- [7] H. Hao, I. Leven, and T. Head-Gordon. "Can electric fields drive chemistry for an aqueous microdroplet?" *Nat. Commun.* 13 (2022), p. 280.
- [8] B. Schirmer and S. Grimme. "Electric field induced activation of H2—Can DFT do the job?" Chem. Commun. 46 (2010), pp. 7942–7944.
- [9] M. Ashton, A. Mishra, J. Neugebauer, and C. Freysoldt. "Ab initio Description of Bond Breaking in Large Electric Fields". Phys. Rev. Lett. 124 (2020), p. 176801.
- [10] M. Stengel, N. A. Spaldin, and D. Vanderbilt. "Electric displacement as the fundamental variable in electronic-structure calculations". Nat. Phys. 5 (2009), pp. 304–308.
- [11] C. Zhang and M. Sprik. "Computing the dielectric constant of liquid water at constant dielectric displacement". Phys. Rev. B 93 (2016), p. 144201.
- [12] T. Sayer, C. Zhang, and M. Sprik. "Charge compensation at the interface between the polar NaCl(111) surface and a NaCl aqueous solution". J. Chem. Phys. 147 (2017), p. 104702.

- [13] S. J. Cox and M. Sprik. "Finite field formalism for bulk electrolyte solutions". J. Chem. Phys. 151 (2019), p. 064506.
- [14] D. Marx and J. Hutter. Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods. Cambridge University Press, Cambridge, 2009.
- [15] N. J. English and C. J. Waldron. "Perspectives on external electric fields in molecular simulation: progress, prospects and challenges". Phys. Chem. Chem. Phys. 17 (2015), pp. 12407–12440.
- [16] H. Elgabarty, N. K. Kaliannan, and T. D. Kühne. "Enhancement of the local asymmetry in the hydrogen bond network of liquid water by an ultrafast electric field pulse". Sci. Rep. 9 (2019), p. 10002.
- [17] H. Elgabarty, T. Kampfrath, D. J. Bonthuis, V. Balos, N. K. Kaliannan, P. Loche, R. R. Netz, M. Wolf, T. D. Kühne, and M. Sajadi. "Energy transfer within the hydrogen bonding network of water following resonant terahertz excitation". Sci. Adv. 6 (2020), eaay7074.
- [18] C. Zhang, T. Sayer, J. Hutter, and M. Sprik. "Modelling electrochemical systems with finite field molecular dynamics". JPhys Energy 2 (2020), p. 032005.
- [19] M. Jia, C. Zhang, and J. Cheng. "Origin of Asymmetric Electric Double Layers at Electrified Oxide/Electrolyte Interfaces". J. Phys. Chem. Lett. 12 (2021), pp. 4616–4622.
- [20] C. Zhang and M. Sprik. "Computing the dielectric constant of liquid water at constant dielectric displacement". Phys. Rev. B 93 (2016), p. 144201.
- [21] Z. Futera and N. J. English. "Water Breakup at Fe2O3-Hematite/Water Interfaces: Influence of External Electric Fields from Nonequilibrium Ab Initio Molecular Dynamics". J. Phys. Chem. Lett. 12 (2021), pp. 6818–6826.
- [22] J. Huang, Y. Zhang, M. Li, A. Groß, and S. Sakong. "Comparing Ab Initio Molecular Dynamics and a Semiclassical Grand Canonical Scheme for the Electric Double Layer of the Pt(111)/Water Interface". J. Phys. Chem. Lett. 14 (2023), pp. 2354–2363.
- [23] J. Behler and M. Parrinello. "Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces". Phys. Rev. Lett. 98 (2007), p. 146401.
- [24] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi. "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons". *Phys. Rev. Lett.* 104 (2010), p. 136403.
- [25] J. Behler. "Four generations of high-dimensional neural network potentials". Chem. Rev. 121 (2021), pp. 10037–10072.
- [26] L. Zhang, M. Chen, X. Wu, H. Wang, W. E, and R. Car. "Deep neural network for the dielectric response of insulators". Phys. Rev. B 102 (2020), p. 041121.
- [27] G. M. Sommers, M. F. Calegari Andrade, L. Zhang, H. Wang, and R. Car. "Raman spectrum and polarizability of liquid water from deep neural networks". Phys. Chem. Chem. Phys. 22 (2020), pp. 10592–10602.
- [28] K. Schütt, O. Unke, and M. Gastegger. "Equivariant message passing for the prediction of tensorial properties and molecular spectra". Proc. Mach. Learn. Res. 2021, pp. 9377–9388.

- [29] D. M. Wilkins, A. Grisafi, Y. Yang, K. U. Lao, R. A. DiStasio, and M. Ceriotti. "Accurate molecular polarizabilities with coupled cluster theory and machine learning". Proc. Natl. Acad. Sci. U.S.A. 116 (2019), pp. 3401–3406.
- [30] V. Kapil, D. M. Wilkins, J. Lan, and M. Ceriotti. "Inexpensive modeling of quantum dynamics using path integral generalized Langevin equation thermostats". J. Chem. Phys. 152 (2020), p. 124104.
- [31] S. Shepherd, J. Lan, D. M. Wilkins, and V. Kapil. "Efficient Quantum Vibrational Spectroscopy of Water with High-Order Path Integrals: From Bulk to Interfaces". J. Phys. Chem. Lett. 12 (2021), pp. 9108–9114.
- [32] R. Beckmann, F. Brieuc, C. Schran, and D. Marx. "Infrared Spectra at Coupled Cluster Accuracy from Neural Network Representations". J. Chem. Theory Comput. 18 (2022), pp. 5492–5501.
- [33] P. Schienbein. "Spectroscopy from machine learning by accurately representing the atomic polar tensor". J. Chem. Theory Comput. 19 (2023), pp. 705–712.
- [34] A. S. Christensen, F. A. Faber, and O. A. von Lilienfeld. "Operators in quantum machine learning: Response properties in chemical space". J. Chem. Phys. 150 (2019), p. 064105.
- [35] M. Gastegger, K. T. Schütt, and K.-R. Müller. "Machine learning of solvent effects on molecular spectra and reactions". Chem. Sci. 12 (2021), pp. 11473–11483.
- [36] A. Gao and R. C. Remsing. "Self-consistent determination of long-range electrostatics in neural network potentials". Nat. Commun. 13 (2022), p. 1572.
- [37] Y. Shao, L. Andersson, L. Knijff, and C. Zhang. "Finite-field coupling via learning the charge response kernel". Electron. Struct. 4 (2022), p. 014012.
- [38] Y. Zhang and B. Jiang. "Universal machine learning for the response of atomistic systems to external fields". Nat. Commun. 14 (2023), p. 6424.
- [39] W. B. Person and J. H. Newton. "Dipole moment derivatives and infrared intensities. I. Polar tensors". J. Chem. Phys. 61 (1974), pp. 1040–1049.
- [40] C. Schran, K. Brezina, and O. Marsalek. "Committee neural network potentials control generalization errors and enable active learning". J. Chem. Phys. 153 (2020), p. 104105.
- [41] A. Szabo and N. S. Ostlund. Modern Quantum Chemistry. Dover Publications, Inc., 1996.
- [42] R. Resta. "Theory of the electric polarization in crystals". Ferroelectrics 136 (1992), pp. 51–55.
- [43] R. Resta and D. Vanderbilt. "Theory of polarization: a modern approach". *Physics of ferroelectrics: a modern perspective*. Springer, pp. 31–68.
- [44] S. K. Natarajan and J. Behler. "Neural network molecular dynamics simulations of solid-liquid interfaces: water at low-index copper surfaces". Phys. Chem. Chem. Phys. 18 (2016), pp. 28704–28725.

- [45] P. Schienbein and J. Blumberger. "Nanosecond solvation dynamics of the hematite/liquid water interface at hybrid DFT accuracy using committee neural network potentials". *Phys. Chem. Chem. Phys.* 24 (2022), pp. 15365–15375.
- [46] C. Schran, F. L. Thiemann, P. Rowe, E. A. Müller, O. Marsalek, and A. Michaelides. "Machine learning potentials for complex aqueous systems made simple". Proc. Natl. Acad. Sci. 118 (2021), e2110077118.
- [47] A. Montenegro, C. Dutta, M. Mammetkuliev, H. Shi, B. Hou, D. Bhattacharyya, B. Zhao, S. B. Cronin, and A. V. Benderskii. "Asymmetric response of interfacial water to applied electric fields". *Nature* 594 (2021), pp. 62–65.
- [48] M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland. "Nuclear Quantum Effects in Water and Aqueous Systems: Experiment, Theory, and Current Challenges". Chem. Rev. 116 (2016), pp. 7529–7550.
- [49] D. P. Fernández, A. R. H. Goodwin, E. W. Lemmon, J. M. H. Levelt Sengers, and R. C. Williams. "A Formulation for the Static Permittivity of Water and Steam at Temperatures from 238 K to 873 K at Pressures up to 1200 MPa, Including Derivatives and Debye-Hückel Coefficients". J. Phys. Chem. Ref. Data 26 (1997), pp. 1125–1166.
- [50] J. G. Kirkwood. "The dielectric polarization of polar liquids". J. Chem. Phys. 7 (1939), pp. 911–919.
- [51] M. Neumann and O. Steinhauser. "On the calculation of the frequency-dependent dielectric constant in computer simulations". Chem. Phys. Lett. 102 (1983), pp. 508–513.
- [52] S. W. de Leeuw, J. W. Perram, E. R. Smith, and J. S. Rowlinson. "Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants". Proc. R. Soc. Lond. A. 373 (1980), pp. 27–56.
- [53] D. Lu, F. m. c. Gygi, and G. Galli. "Dielectric Properties of Ice and Liquid Water from First-Principles Calculations". Phys. Rev. Lett. 100 (2008), p. 147601.
- [54] T. Morawietz, A. Singraber, C. Dellago, and J. Behler. "How van der Waals interactions determine the unique properties of water". Proc. Natl. Acad. Sci. 113 (2016), pp. 8368– 8373.
- [55] M. Neumann and O. Steinhauser. "Computer simulation and the dielectric constant of polarizable polar systems". Chem. Phys. Lett. 106 (1984), pp. 563–569.
- [56] J. E. Bertie and Z. Lan. "Infrared Intensities of Liquids XX: The Intensity of the OH Stretching Band of Liquid Water Revisited, and the Best Current Values of the Optical Constants of H₂O(l) at 25 °C between 15 000 and 1 cm⁻¹". Appl. Spectrosc. 50 (1996), pp. 1047–1057.
- [57] R. Rey, K. B. Møller, and J. T. Hynes. "Hydrogen Bond Dynamics in Water and Ultrafast Infrared Spectroscopy". J. Phys. Chem. A 106 (2002), pp. 11993–11996.
- [58] C. P. Lawrence and J. L. Skinner. "Vibrational spectroscopy of HOD in liquid D₂O. III. Spectral diffusion, and hydrogen-bonding and rotational dynamics". J. Chem. Phys. 118 (2003), pp. 264–272.

- [59] C. J. Fecko, J. D. Eaves, J. J. Loparo, A. Tokmakoff, and P. L. Geissler. "Ultrafast Hydrogen-Bond Dynamics in the Infrared Spectroscopy of Water". Science 301 (2003), pp. 1698–1702.
- [60] P. Schienbein and D. Marx. "Supercritical Water is not Hydrogen Bonded". Angew. Chem. Int. Ed. 59 (2020), pp. 18578–18585.
- [61] G. Cassone, J. Sponer, S. Trusso, and F. Saija. "Ab initio spectroscopy of water under electric fields". Phys. Chem. Chem. Phys. 21 (2019), pp. 21205–21212.
- [62] J. D. Jackson. Classical electrodynamics. 3rd ed. New York, NY: Wiley, 1999.
- [63] M. Cools-Ceuppens, J. Dambre, and T. Verstraelen. "Modeling Electronic Response Properties with an Explicit-Electron Machine Learning Potential". J. Chem. Theory Comput. 18 (2022), pp. 1672–1691.
- [64] P. Schienbein and D. Marx. "Liquid-Vapor Phase Diagram of RPBE-D3 Water: Electronic Properties along the Coexistence Curve and in the Supercritical Phase". J. Phys. Chem. B 122 (2018), pp. 3318–3329.
- [65] P. Schienbein and D. Marx. "Assessing the properties of supercritical water in terms of structural dynamics and electronic polarization effects". Phys. Chem. Chem. Phys. 22 (2020), pp. 10462–10479.
- [66] S. Imoto, H. Forbert, and D. Marx. "Water structure and solvation of osmolytes at high hydrostatic pressure: pure water and TMAO solutions at 10 kbar versus 1 bar". Phys. Chem. Chem. Phys. 17 (2015), pp. 24224–24237.
- [67] K. Forster-Tonigold and A. Groß. "Dispersion Corrected RPBE Studies of Liquid Water". J. Chem. Phys. 141 (2014), p. 064501.
- [68] A. Groß and S. Sakong. "Ab Initio Simulations of Water/Metal Interfaces". Chem. Rev. 122 (2022), pp. 10746–10776.
- [69] M. Geiger, T. Smidt, A. M., B. K. Miller, W. Boomsma, B. Dice, K. Lapchevskyi, M. Weiler, M. Tyszkiewicz, S. Batzner, D. Madisetti, M. Uhrin, J. Frellsen, N. Jung, S. Sanborn, M. Wen, J. Rackers, M. Rød, and M. Bailey. *Euclidean neural networks: e3nn.* Version 0.5.0. 2022.
- [70] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". Adv. Neural. Inf. Process. Syst. 32. Curran Associates, Inc., 2019, pp. 8024–8035.
- [71] V. Kapil, M. Rossi, O. Marsalek, R. Petraglia, Y. Litman, T. Spura, B. Cheng, A. Cuzzocrea, R. H. Meißner, D. M. Wilkins, et al. "i-PI 2.0: A universal force engine for advanced molecular simulations". Comput. Phys. Commun. 236 (2019), pp. 214–223.
- [72] A. Singraber, T. Morawietz, J. Behler, and C. Dellago. "Parallel multistream training of high-dimensional neural network potentials". J. Chem. Theory Comput. 15 (2019), pp. 3075–3092.

- [73] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, and J. Hutter. "CP2K: An electronic structure and molecular dynamics software package Quickstep: Efficient and accurate electronic structure calculations". J. Chem. Phys. 152 (2020), p. 194103.
- [74] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter. "Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach". Comput. Phys. Commun. 167 (2005), pp. 103–128.
- [75] M. A. Marques, M. J. Oliveira, and T. Burnus. "Libxc: A Library of Exchange and Correlation Functionals for Density Functional Theory". Comput. Phys. Commun. 183 (2012), pp. 2272–2281.
- [76] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu". J. Chem. Phys. 132 (2010), p. 154104.
- [77] G. Lippert, J. Hutter, and M. Parrinello. "A Hybrid Gaussian and Plane Wave Density Functional Scheme". Mol. Phys. 92 (1997), pp. 477–488.
- [78] J. VandeVondele and J. Hutter. "Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases". J. Chem. Phys. 127 (2007), p. 114105.
- [79] S. Goedecker, M. Teter, and J. Hutter. "Separable dual-space Gaussian pseudopotentials". Phys. Rev. B 54 (1996), pp. 1703–1710.
- [80] C. Hartwigsen, S. Goedecker, and J. Hutter. "Relativistic separable dual-space Gaussian pseudopotentials from H to Rn". *Phys. Rev. B* 58 (1998), pp. 3641–3662.
- [81] P. Umari and A. Pasquarello. "Ab initio Molecular Dynamics in a Finite Homogeneous Electric Field". *Phys. Rev. Lett.* 89 (2002), p. 157602.
- [82] R. M. Pick, M. H. Cohen, and R. M. Martin. "Microscopic Theory of Force Constants in the Adiabatic Approximation". Phys. Rev. B 1 (1970), pp. 910–920.
- [83] E. Ditler, C. Kumar, and S. Luber. "Analytic calculation and analysis of atomic polar tensors for molecules and materials using the Gaussian and plane waves approach". The Journal of Chemical Physics 154 (2021), p. 104121.
- [84] V. P. Nicu, J. Neugebauer, S. K. Wolff, and E. J. Baerends. "A vibrational circular dichroism implementation within a Slater-type-orbital based density functional framework and its application to hexa-and hepta-helicenes". Theoretical Chemistry Accounts 119 (2008), pp. 245–263.
- [85] O. T. Unke and M. Meuwly. "PhysNet: A neural network for predicting energies, forces, dipole moments, and partial charges". *Journal of chemical theory and computation* 15 (2019), pp. 3678–3693.

- [86] O. T. Unke, S. Chmiela, M. Gastegger, K. T. Schütt, H. E. Sauceda, and K.-R. Müller. "SpookyNet: Learning force fields with electronic degrees of freedom and nonlocal effects". Nature communications 12 (2021), p. 7273.
- [87] Z. Song, J. Han, G. Henkelman, and L. Li. "Charge-Optimized Electrostatic Interaction Atom-Centered Neural Network Algorithm". *Journal of Chemical Theory and Computation* (2024).
- [88] G. Bussi, D. Donadio, and M. Parrinello. "Canonical sampling through velocity rescaling". J. Chem. Phys. 126 (2007).
- [89] S. Nosé. "A unified formulation of the constant temperature molecular dynamics methods". J. Chem. Phys. 81 (1984), pp. 511–519.
- [90] S. Nosé. "A molecular dynamics method for simulations in the canonical ensemble". Mol. Phys. 52 (1984), pp. 255–268.

Acknowledgements

K.J. gratefully acknowledges a PhD studentship co-sponsored by University College London and Pacific Northwest National Laboratory (PNNL) through its BES Geosciences program supported by the U.S. Department of Energy's Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences and Biosciences Division. This work was further supported by an individual postdoc grant to P.S. funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project number 519139248 (Walter Benjamin Programme). J.B. and P.S acknowledge EPSRC - UKRI for award of computing time via a ARCHER2 Pioneer Project (ARCHER2 PR17125). Via our membership of the UK's HEC Materials Chemistry Consortium, which is funded by EPSRC (EP/L000202, EP/R029431), this work also used the ARCHER UK National Supercomputing Service (http://www.archer.ac.uk).

Author Contributions

K.J. implemented PNNP under the supervision of P.S., K.J. carried out and analyzed all simulations. P.S. calculated the IR spectra and helped analyzing the data. P.S and J.B designed the research. K.J., P.S. and J.B wrote the manuscript, with input from K.R. All authors reviewed and discussed the manuscript.

Competing Interests

The authors declare no competing interests.

Additional Information

Supplementary information. The online version contains supplementary material available.

Supporting Information

Molecular dynamics simulation with finite electric fields using Perturbed Neural Network Potentials

Kit Joll¹, Philipp Schienbein *1,2 , Kevin M. Rosso³, and Jochen Blumberger $^{\dagger 1}$

¹Department of Physics and Astronomy and Thomas Young Centre, University College London, London, WC1E 6BT, United Kingdom

²Department of Physics, Imperial College London, Exhibition Rd, South Kensington, London, SW7 2AZ, United Kingdom

 $^3 \mbox{Pacific Northwest National Laboratory, Richland, Washington 99354, United States$

 $^{^*\}mbox{to}$ whom correspondence should be addressed, email: p.schienbein@ucl.ac.uk

[†]to whom correspondence should be addressed, email: j.blumberger@ucl.ac.uk

Energy, Charge and Momentum conservation

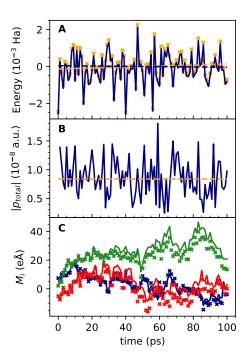


Figure S1: Panel A shows the difference between the field-induced energy contribution and its mean value as a function of simulation time calculated from the integrated polarization (solid blue line) and from DFT reference calculations (yellow crosses). The dashed lines illustrate linear fits to that data using the same color code. Panel B shows the magnitude of the total momentum as a function of simulation time. Panel C presents the polarization as a function of simulation time, calculated by integration (Eq. 5, solid lines) and from explicit DFT reference calculations (crosses) for the three spatial coordinatex x (blue), y (red), and z (green). Mind that the electric field is applied along the z direction.

Field Sweep

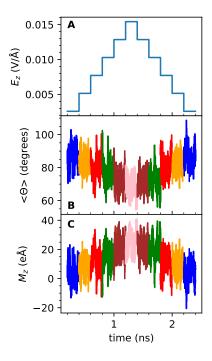


Figure S2: PNNP MD simulating a field sweep going from 0 to $0.0154\,\mathrm{V\,\mathring{A}}^{-1}$, by increasing the applied field strengths by $0.0026\,\mathrm{V\,\mathring{A}}^{-1}$ every 200 ps. The applied field strength is depicted in panel A, while the response of the water molecules in terms of the average water orientation $\langle\Theta\rangle$ (see main text) and of the polarization along the field (M_z) is shown in panels B and C, respectively.

Convergence of dielectric constant

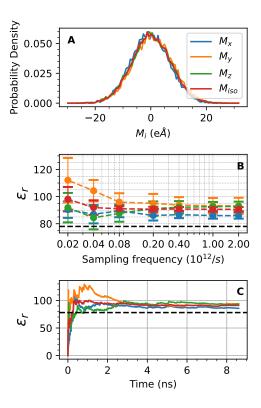


Figure S3: Convergence of the dielectric constant as a function of time using the variance of the total dipole moment along a zero field MD simulation, see Eq. 8 in the main text. Panel A depicts the ensemble distribution function of the total dipole moment individually for the different cartesian components and their average ($M_{\rm iso}$, red). The convergence of the dielectric constant is presented as a function of the sampling frequency (panel B) and the total simulation time (C), both following the same color code as in panel A.

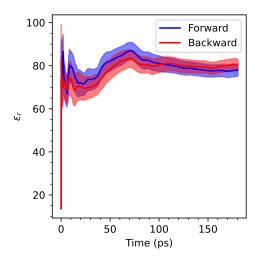


Figure S4: Convergence of the dielectric constant as a function of time during the forward (blue) and backward (red) electric sweep PNNP MD simulation depicted in Fig. S2 using the change in polarization caused by an applied electric field Eq. 7. The shaded area illustrates the corresponding standard deviation of the mean.